PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C12Q 1/68, C07H 21/04	A1	(11) International Publication Number: WO 91/10745
		(43) International Publication Date: 25 July 1991 (25.07.91)

(21) International Application Number: PCT/GB91/00066

(22) International Filing Date: 17 January 1991 (17.01.91)

(30) Priority data:
9001181.8 18 January 1990 (18.01.90) GB

(71) Applicant *(for all designated States except US)*: IMPERIAL CANCER RESEARCH TECHNOLOGY LIMITED [GB/GB]; Sardinia House, Sardinia Street, London WC2A 3NL (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)* : WOLF, Charles, Roland [GB/GB]; Imperial Cancer Research Fund, Laboratory of Molecular Pharmacology and Drug Metabolism, Department of Biochemistry, Hugh Robson Building, George Square, Edinburgh EH8 9XD (GB). MILES, John, Stephen [GB/GB]; Department of Biochemistry, The University of Glasgow, Glasgow G12 8QQ (GB). SPURR, Nigel, Kay [GB/GB]; GOUGH, Alan, Charles [GB/GB]; Imperial Cancer Research Fund, Clare Hall Laboratories, Blanche Lane, South Mimms, Potters Bar, Herts EN6 3LD (GB).

(74) Agent: BASSETT, Richard, S.; Eric Potter & Clarkson, St Mary's Court, St Mary's Gate, Nottingham NG1 1LE (GB).

(81) Designated States: AT (European patent), AU, BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB, GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent), US.

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: GENETIC ASSAY FOR CYTOCHROME P450

(57) Abstract

Polymorphisms at positions 100, 271, 281, 294, 408, 506 or 1432 of the cytochrome P450 enzyme bufuralol-1'-hydroxylase are indicative of the extensive metaboliser/poor metaboliser phenotypes and can be detected using known methods such as amplification of the DNA with the polymerase chain reaction, followed by digestion with a suitable restriction enzyme and analysis by gel electrophoresis. The EM/PM phenotype is relevant to calculating safe or effective drug doses for individuals.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**ES**	Spain	**MG**	Madagascar
AU	Australia	**FI**	Finland	**ML**	Mali
BB	Barbados	**FR**	France	**MN**	Mongolia
BE	Belgium	**GA**	Gabon	**MR**	Mauritania
BF	Burkina Faso	**GB**	United Kingdom	**MW**	Malawi
BG	Bulgaria	**GN**	Guinea	**NL**	Netherlands
BJ	Benin	**GR**	Greece	**NO**	Norway
BR	Brazil	**HU**	Hungary	**PL**	Poland
CA	Canada	**IT**	Italy	**RO**	Romania
CF	Central African Republic	**JP**	Japan	**SD**	Sudan
CG	Congo	**KP**	Democratic People's Republic of Korea	**SE**	Sweden
CH	Switzerland			**SN**	Senegal
CI	Côte d'Ivoire	**KR**	Republic of Korea	**SU**	Soviet Union
CM	Cameroon	**LI**	Liechtenstein	**TD**	Chad
CS	Czechoslovakia	**LK**	Sri Lanka	**TG**	Togo
DE	Germany	**LU**	Luxembourg	**US**	United States of America
DK	Denmark	**MC**	Monaco		

Genetic assay for cytochrome P 450

The present invention relates to a genetic assay, that is to say an assay which reveals the presence or absence of a genetic characteristic.

It is known that mutations in regions of the nucleic acids of organisms can alter the nature or amount of polypeptides encoded by such regions or encoded by other regions associated with the site of mutation.

We have now found several sites of mutation in mammalian DNA associated with the cytochrome P450-dependent monooxygenase supergene family of enzymes. The presence or absence of mutation at one or more of these sites has been found to indicate with a high degree of certainty whether the individual is an "extensive metaboliser" or a "poor metaboliser".

Skoda *et al* (1988 *P.N.A.S.* 85, 5240-5243) disclosed an RFLP-based assay which identifies only about 25% of poor metabolisers.

Cytochrome P450-dependent monooxygenases (P450s) are a supergene family of enzymes that catalyse the oxidation of lipophilic chemicals through the insertion of one atom of molecular oxygen into the substrate. They are involved in the metabolism of xenobiotic compounds, and in particular with the

SUBSTITUTE SHEET
ISA/EP

clearance of at least 25 drugs, including debrisoquine, sparteine, bufuralol and dextromethorphan. Other drugs whose metabolism is related to the debrisoquine oxidation polymorphism (as of June 1990) include (in the cardiovascular area) metoprolol, timolol, propranolol, perhexilene, N-propylamaline, propafenone, encainide, flecainide and mexiletine, (in the psychiatric area) amitriptyline, imipramine, desipramine, nortriptyline, clomipramine, thioridazine, perphenazine, amiflamine and tomoxitene and (in other areas) codeine, methoxyphenamine and phenformin and possibly also chlorpropamine, melatonin and MPTP. The P450 system is polymorphic in man, and genetic differences in the P450-mediated metabolism of a wide variety of drugs have been clearly demonstrated. The best example of this is the debrisoquine/sparteine polymorphism, (see Ref 1 for a review). Up to 10% of the Caucasian population exhibit the poor metaboliser (PM) phenotype. This is characterised by a significantly reduced ability to metabolise the prototype drug debrisoquine to 4-hydroxydebrisoquine, the metabolism being 10-200 times less than in extensive metabolisers (EMs). The PM phenotype is inherited as an autosomal recessive trait, and up to 54% of people are carriers of a mutant allele(s). The PM phenotype leads to impaired clearance of over twenty other commonly prescribed drugs and may result in serious adverse side effects upon their administration. Thus the ability to predict phenotype is an attractive possibility which would be useful in

SUBSTITUTE SHEET
ISA/EP

many clinical situations.

Recently the cytochrome P450 isozyme (P450 db1, also called P450 buf 1 or P450 DB) responsible for the metabolism of debrisoquine, sparteine and other compounds related to the polymorphism has been purified from human liver. Immunoquantitation of this protein correlates well with the levels of bufuralol-1'-hydroxylase activity in a series of human livers, bufuralol being a prototype substrate for the db1 isozyme. Furthermore no immuno-reactive db1 protein was detected in liver microsomes of PMs suggesting that the complete or almost complete absence of this protein leads to the PM phenotype. Recent work also provides evidence for the presence of allozymes of P450db1 with altered K_m and V_{max} probably due to amino acid substitutions. Antibodies against human P450db1 have been found in patients with autoimmune hepatitis type II but the relationship between the debrisoquine polymorphism and the appearance of these autoantibodies is not known.

Gonzalez and coworkers have isolated cDNA clones from libraries made from the livers of EMs and have shown that they encode active P450db1 by expression in COS-1 cells and measurement of bufuralol-1'-hydroxylase activity (2). Sequence analysis shows that P450 db1 belongs to a distinct P450 subfamily, P450IID (3). P450IID cDNA clones were also obtained from libraries made from the livers of PMs and in these cases

SUBSTITUTE SHEET
ISA/EP

they appeared to be derived from aberrantly or incompletely spliced mRNAs, and therefore would not be able to encode an active P450db1. Four variants were described: "a" which retains intron 5; "b" which retains intron 6; "b'" which has lost the 3' half of exon 6 in combination with the removal of intron 6; and another cDNA clone from a PM liver, variant "c", which appears to be normally spliced but has several base substitutions and was not characterised further. It was inferred from these studies that the defective mRNAs (cDNAs) were the products of mutant alleles of the P450db1 gene.

The gene encoding P450db1 (*CYP2D1*) has been located on chromosome 22 and Southern blot analysis shows that there is probably more than one gene/pseudogene within the *CYP2D* locus based on the amount of DNA hybridizing to the db1 cDNA probe. The *CYP2D* locus is highly polymorphic, and two alleles, detected using the restriction enzyme *Xba*I, have been associated with the PM phenotype (44kb allele and 11.5kb allele; 4). However, at the present time these restriction fragment length polymorphisms (RFLPs) are not informative in predicting phenotype as they do not identify all PM individuals (4).

SUBSTITUTE SHEET
ISA/EP

We have now cloned and sequenced further novel P450IID cDNAs, none of which we predict would encode an active P450. By comparison with the available P450IID cDNA sequences, we have identified base-pair differences which form the basis of a genotyping assay for the PM phenotype.

One aspect of the invention provides a method of detecting a mutation at positions 100, 271, 281, 294, 408, 506 or 1432 of the DNA sequence of P450IID bufuralol-1'-hydroxylase or a deletion of at least part of exon 9 thereof.

The enzyme is also called debrisoquine hydroxylase and P450IID6 and is encoded by gene *CYPD2D6*.

In the case of positions 100, 271, 281, 294, 408 and 1432, the mutation is typically one or more base pair substitutions such as C -> T, C -> A, A -> G, C -> G, G -> C or C -> T respectively. (These alterations are more fully written as cytosine to thymidine, cytosine to adenosine, adenosine to guanine, cytosine to guanine, guanine to cytosine and cytosine to thymidine.) In the case of a polymorphism at position 506, the mutation is typically a base pair deletion in the spliced product, resulting, apparently, from a G -> A transition in the last nucleotide of intron 3. Since the assay will normally be directed to genomic DNA, it is the G -> A transition which is detected directly.

The mutations at positions 100, 271, 281 and 294 are either silent or lead to single amino acid substitutions. In themselves, they do not account for the PM phenotype but they are strongly linked with the base pair deletion at 506 and are therefore informative.

The assay may involve any suitable method for identifying such polymorphisms, such as: sequencing of the DNA at one or more of the relevant positions; differential hybridisation of an oligonucleotide probe designed to hybridise at the relevant positions of either the wild-type or mutant sequence; denaturing gel electrophoresis following digestion with an appropriate restriction enzyme, preferably following amplification of the relevant DNA regions; S1 nuclease sequence analysis; non-denaturing gel electrophoresis, preferably following amplification of the relevant DNA regions; conventional RFLP (restriction fragment length polymorphism) assays; selective DNA amplification using oligonucleotides which are matched for the EM sequence and unmatched for the PM sequence or *vice versa*; or the selective introduction of a restriction site using a PCR (or similar) primer matched for the PM or EM genotype, followed by a restriction digest. The assay may be indirect, ie capable of detecting a mutation at another position or gene which is known to be linked to one or more of the positions listed above, especially the deletion at position 506. Assays directed to the related locus *CYP2D7* may be used in this way. A number of sites

SUBSTITUTE SHEET
ISA/EP

in the "b" variant sequence (Gonzalez *et al*) have recently been analysed and also been shown to be informative for the PM phenotype. The "b" variant appears to be derived from a gene other than the P450IID6 but its presence is tightly linked to the PM phenotype and the presence of the "a" variant. Base pair differences between the "b" variant and other genes in this cluster which may be used as a basis for a genotyping assay include for example, bp 632 (G insertion); bp 637 and 638 (TC to CT); bp 691 (C to T); bp 832 (A to G); bp 1085 (T insertion); 1094 (G to A); bp 1528 (T to C). The probes and primers may be fragments of DNA isolated from nature or may be synthetic.

A non-denaturing gel may be used to detect differing lengths of fragments resulting from digestion with an appropriate restriction enzyme. The DNA is usually amplified before digestion, for example using the polymerase chain reaction (PCR) method disclosed in reference 5 and modifications thereof. Otherwise 10-100 times more DNA would need to be obtained in the first place, and even then the assay would work only if the restriction enzyme cuts DNA infrequently.

Amplification of DNA may be achieved by the established PCR method or by developments thereof or alternatives such as the ligase chain reaction, QB replicase and nucleic acid sequence-based amplification.

SUBSTITUTE SHEET
ISA/EP

An "appropriate restriction enzyme" is one which will recognise and cut the wild-type sequence and not the mutated sequence or *vice versa*. The sequence which is recognised and cut by the restriction enzyme (or not, as the case may be) can be present as a consequence of the mutation or it can be introduced into the normal or mutant allele using mismatched oligonucleotides in the PCR reaction. Various enzymes are disclosed below as specific examples, but any enzyme which cuts at the same place (an "isoschizomer") or which recognises the same sequence and cuts the DNA at a point within or adjacent the sequence will be suitable; more are being discovered all the time. It is convenient if the enzyme cuts DNA only infrequently, in other words if it recognises a sequence which occurs only rarely.

Restriction enzymes useful in connection with the mutations described above include, for example, *Hae*III for position 294, *Sac*II for 281, *Hha*I for 271, *Eco*RI for 408, *Bst*NI for 506 and *Apy*I for 100. These enzymes are available commercially from suppliers of biological reagents, such as BRL-Gibco, Paisley, Scotland.

In another method, a pair of PCR primers are used which match (ie hybridise to) either the PM genotype or the EM genotype but not both. Whether amplified DNA is produced will then indicate the PM or EM genotype (and hence phenotype).

SUBSTITUTE SHEET
ISA/EP

However, this method relies partly on a negative result (ie the absence of amplified DNA) which could be due to a technical failure. It is therefore less reliable and/or requires additional control experiments.

A preferable method employs similar PCR primers but, as well as hybridising to only one of the PM or EM sequences, they introduce a restriction site which is not otherwise there in either the PM or EM sequences. For example, PCR primers G and H:

(G) 5'-GATGAGCTGCTAACTGAGCCC-3'

(H) 5'-CCGAGAGCATACTCGGGAC-3'

will introduce a *Msp*I site at the 775 region in the PM sequence. Neither the PM nor EM sequences have a *Msp*I site at that position. Thus, in a single two-step process of PCR amplification with primers D, E, G and H followed by a restriction digest with *Bst*I and *Msp*I, both mutations can be detected.

The nucleic acid, usually genomic DNA rather than RNA, which is assayed may be obtained from any cell of the body (such as hair roots, buccal epithelial cells and blood spots) or even urinary deposits. Conveniently, a mouthwash or drop of blood is taken, either of which will contain a few cells. Preferably, the

SUBSTITUTE SHEET
ISA/EP

DNA is extracted by known techniques and a specific region of the P450IID sequence is amplified using the PCR. It is then digested with the restriction enzyme and subjected to PAGE (polyacrylamide gel electrophoresis). The gel is stained and photographed to reveal a pattern of fragments indicative of whether the patient is homozygous EM/EM, homozygous PM/PM or heterozygous. The whole procedure, using current technology, takes about 5 hours whereas existing methods in which drug metabolism is monitored take up to 3 days and are much more difficult to perform.

A kit may be provided, according to another aspect of the invention, to perform the assay. The kit will typically contain the primer(s) needed for the PCR amplification (if PCR amplification is used) and also control DNA for both homozygotes and the heterozygote, so that the results of the assay can be analysed more readily. Conveniently, the kit also comprises the restriction enzyme(s) and, preferably, phenol and SDS (sodium dodecyl sulphate) or similar materials used in the mouthwash.

The assays of the invention will be extremely valuable in relation to human medicine and may be used prior to treatment with a drug which is toxic if not metabolised or which is effective only if metabolised. They may also be used to identify individuals who are genetically predisposed to be susceptible to or resistant to diseases the etiology of which is

SUBSTITUTE SHEET
ISA/EP

linked to the PM/EM phenotype, for example lung and bladder cancer. The mutations described above are known to occur in European Caucasians and may or may not be present in other races such as Mongoloids and Negroids.

It is entirely possible that other mutations will be found which indicate the PM phenotype. If so, the assays of the invention may be used together with assays for such other mutations in order to provide a definitive PM/EM phenotyping assay.

The assays of the invention may be used as part of the clinical trials of a new drug: by phenotyping the healthy volunteers or patients in the trials and conducting appropriate drug metabolism studies, it can be established whether the drug's metabolism is affected by the PM/EM phenotype.

Examples of the invention will now be described with reference to the accompanying drawings, the legends for which are as follows:

Figure 1: This is a diagrammatic representation of the db1-related cDNA clones aligned with db1. Clones pMP32 and pMP33 are from this work, and db1 and "a" taken from Gonzalez *et al* (2). The numbering refers to that for db1 starting at the initiation codon. The positions of the introns, inferred from

SUBSTITUTE SHEET
ISA/EP

comparison with other P450II genes, are indicated on db1 by vertical lines, and the retention of part of intron 1 (pMP33) and intron 5 ("a") is indicated by triangles. Base pair substitutions and other differences between the cDNA sequences are given in full in Fig. 2. The dashes represent sequence not present in the variant cDNAs compared to db1.

Figure 2 (on two sheets: Figure 2A and Figure 2B): This shows a comparison of nucleotide sequences of DNA encoding functional debrisoquine hydroxylase (db1) and related cDNAs. The sequences for db1 and variant "b" are taken from Gonzalez *et al* (2). Numbering starts at the initiation codon for db1. The variant "a" cDNA sequence was compiled from two cDNAs, pMP32 and pMP33 which contained an identical sequence over an overlapping region of 462 bp. This sequence contained only one base-pair difference to the partial sequence described as variant "a" by Gonzalez *et al* (G to C at position 383) which covers the region 299-1567 bp. The base-pair deletion at position 506 is marked *. The position of introns (vertical lines) is obtained from the sequence of P45011D6 gene (11). Triangles represent the positions of intron sequences in the isolated variant cDNAs:

1. The first 64-bp of intron 1 in pMP33; 5. Insertion of intron 5 in the variant "a" cDNA described by Gonzalez *et al* but which was absent from pMP32 and pMP33; 6. Insertion of part of intron 6 in both pMP34 and variant "b". The position of the bp insertions (G) at position 631 and at 983 (T) in the "b" variant

SUBSTITUTE SHEET
ISA/EP

are shown **. The positions of oligonucleotide primers used to amplify specific regions of the genes are marked by capital letters and horizontal arrows. The oligonucleotide marked D was taken from the sequence of intron 4 (see Fig 5). The diagnostic restriction sites used to differentiate between either db1 (variant "a") and variant "b" or db1 and variant "a" are marked.

Figure 3: This is a schematic illustration of *Hae*III cleavage sites in the PCR fragment generated from the db1 gene using oligonucleotides A and B as primers (see Fig 2). The location of the determinative *Hae*III site when position 294 is mutated is shown. The fragment generated is 172bp in length and incorporates the sequences of intron 1 and exon 2.

Figure 4: This is a schematic illustration of the DNA fragmentation pattern obtained following polyacrylamide gel electrophoresis of the *Hae*III-cleaved exon 2 DNA shown in Figure 3 for both homozygote EM, homozygote PM and heterozygote individuals.

Figures 5 & 6: Analysis of the mutation site leading to the base-pair deletion in db1 cDNA at position 506. Diagrammatic representation of the method used and (Figure 6) predicted banding pattern for individuals containing the db1 gene, variant "a" or both. Amplification of DNA from exon 3 into intron 4 produced a 334-bp fragment which in individuals containing the

SUBSTITUTE SHEET
ISA/EP

db1 sequence digests into fragments of 105 bp and 229 bp with *Bst*N1. The variant "a" sequence (PM's) is resistant to digestion with this enzyme.

Figure 7: This shows the "b" variant sequence and others alongside the db1 sequence, with the region around exon 7 shown in more detail in Figure 8. Sites cut by restriction enzymes *Hae*III, *Dra*III and *Hha*I are shown as A, B and C respectively. *Hae*III cuts db1 only. *Dra*III cuts 2D7 (pMP34) and 2d8 sequences only at 1083. A combination of *Hae*III and *Dra*III is useful but inconvenient. A *Hha*I site is absent in the "b" variant at 1094 but present in all other variants so far found. Primer pair 1+2 is used for *Hae*III/*Dra*III analysis and 3+2 for *Hha*I/*Dra*III analysis.

MATERIALS & METHODS

Preparation of radioactive probes. The human P450 db1 (P450IID1; 2) cDNA probe was kindly provided by Drs F.J. Gonzalez and U.A. Meyer. Restriction fragments for use as probes were isolated from low gelling temperature agarose and radioactively labelled with [α-^{32}P]dCTP (3000 Ci mmol^{-1}; Amersham International) by random primer extension (6). Oligonucleotides were made on an Applied Biosystems 380A machine and labelled with [α-^{32}P]ATP (3000 Ci mmol^{-1}; Amersham International) and T4 polynucleotide kinase.

SUBSTITUTE SHEET
ISA/EP

Isolation of human P450IID-related cDNA clones. The full length P450IID6 (db1) cDNA was used to screen plaques from two human liver lambda gt11 cDNA libraries (Clontech, Palo Alto, CA; and Kwok *et al* [7]) made from individuals of unknown phenotype. Two different cDNA clones, lambda MPA (1.22kb) and lambda MPG (1.56kb) from the Clontech library were subcloned into pUC18 to give pMP32 and pMP33 respectively, and also into M13mp18 for sequence analysis.

DNA sequence analysis. The dideoxy chain termination method was used with [α-^{35}S] thio dATP (400 Ci mmol^{-1}; Amersham International) to sequence DNA cloned in M13 (8, 9). Overlapping sequences were derived using a series of synthetic oligo-nucleotides and both DNA strands were fully sequenced. Sequences were compiled and analysed using Staden Plus software implemented on a DCS286 computer (10). DNA sequences generated during the course of this work have been deposited in the EMBL Data Bank with accession nos. X16865 and X16866.

Phenotyping. Individuals GT, MJ, TR, PJ and ML were phenotyped *in vivo* by Prof. G. Tucker, Department of Pharmacology and Therapeutics, University of Sheffield, U.K. using debrisoquine; and individuals A, 1.1, 2.2, and 3.3 were phenotyped using sparteine *in vivo* by us. Post-mortem liver samples E4, E6, E8 and E11 were phenotyped *in vitro* by Dr. U.A. Meyer, Biozentrum,

SUBSTITUTE SHEET
ISA/EP

Basel using bufuralol, and LVII showed no cross-reactivity with a monoclonal antibody raised against rat P450dbi (kindly supplied by Dr Meyer).

DNA amplification. Total DNA was isolated from blood lymphocytes or from liver using an Applied Biosystems 340A nucleic acids extractor according to the manufacturer's instructions. Target DNA (1µg for genomic DNA or 1ng for cloned cDNA) was used in the polymerase chain reaction (PCR; 5) with 600ng of each amplification primer. The PCR was carried out using 2.5U *Taq* DNA polymerase (Cetus Corporation) according to the manufacturer's conditions except that dimethyl sulphoxide was added to 10% (v/v) final concentration. The chain reaction was initiated by denaturing DNA at 92°C for 1 min, annealing by cooling to 60°C for 1 min and extending at 72°C for 2 min; twenty cycles were performed using either a Cetus or Techne programmable heating block. Pairs of oligonucleotide primers enabling the amplification of specific exon sequences were used as described in the legend to Figure 2.

Analysis of amplified DNA. The products of DNA amplification (between 1/20 - 1/10 of the total) were either left uncut or digested with diagnostic restriction enzymes and separated electrophoretically on 6% polyacrylamide gels. In some cases the DNA was analysed by Southern blotting. Briefly, the DNA was transferred to Hybond N by alkali transfer (1.5M NaCl, 0.25M

SUBSTITUTE SHEET
ISA/EP

NaOH) and baked at 80°C for 2h. Membranes to be probed with cDNAs were pre-hybridised at 65°C in 5XSSC, 4X Denhardt's, 10% SDS, 0.1% NaPPi, 20μg ml^{-1} salmon sperm DNA. Hybridisation was overnight in the same buffer except no salmon sperm DNA was present. Filters were washed to a final stringency of 0.2X SSC, 0.1% SDS, 0.1% NaPPi at 65°C. Oligonucleotide probes covering the region of interest were hybridised to the membranes at 37°C in 6X SSC, 0.1% NaPPi and washed in the same buffer at a temperature dependent on their T_m. Membranes were exposed to Kodak XAR-5 film for between 2h and three days at -80°C.

EXAMPLE 1: Comparison of db1-related DNA sequences

The nucleotide sequences derived from pMP32 and pMP33 were compared with the sequences of the db1 cDNA (encoding a functional P450 with bufuralol 1'-hydroxylase activity) as well as the variant "a" cDNA described by Gonzalez *et al* (2) and amended in the EMBL Data Bank; accession no. Y00300 (Figs. 1 & 2). The sequences of pMP32 and pMP33 have an identical sequence over an overlapping region of 445bp. Together these clones constitute a full length cDNA. However, we did not expect the sequence generated by pMP32/pMP33 to encode a functional protein as it contains part of intron 1 and perhaps more importantly a single base deletion at position 506 which would lead to a frame-shift (Fig. 2). Variant "a" of Gonzalez *et al* (2) also contains this same frame-shift and so with or without the retention of intron 5 or intron 1 this variant would not encode a functional protein (Fig 2). Analysis of the genomic P450II2D sequence showed that the bp deletion at position 506 is due to a G to A transition at the junction of intron 3 and exon 4. This mutation removes a *Bst*N1 site in the gene compared to db1.

SUBSTITUTE SHEET
ISA/EP

There are a significant number of base pair differences between the sequence of pMP32/pMP33, representing a full length variant, non-functional db1 sequence, and the normal db1 sequence. All of these differences may serve as markers for the PM phenotype and may therefore be of use in a genotyping assay.

EXAMPLE 2: Analysis at position 294

An analysis at position 294 is shown as an example in Figs 3 and 4. DNA covering the region of interest was amplified by PCR from genomic DNA using the oligonucleotides A and B (Fig 2). This generates a DNA fragment of 172bp in length (Fig 3). In PM individuals the 83bp fragment generated by digestion with the restriction enzyme *Hae*III will cut into two pieces (39 and 44bp) due to the *Hae*III site generated by the mutation at position 294. In homozygous EM's this site is absent. 100% agreement between phenotype, assessed by measuring the rate of metabolism and marker drugs, and genotype was observed in the 16 individuals tested, with the gels corresponding to the expected appearance shown in Figure 4.

Analysis of a region of exon 9 using oligonucleotide primers G and F (see Fig 2) showed that in some PM individuals no dbl related band could be observed, indicating a deletion of this region of the gene. This also is informative for the PM phenotype in some individuals.

EXAMPLE 3: Analysis at position 506

Genomic DNA from individuals phenotyped either as poor or normal metabolizers was amplified by PCR using an oligonucleotide derived from exon 3 and one derived from intron 4 (marked C and D in Fig. 2) using an annealing temperature of 60°C. The sequence of oligonucleotide D was AAATCCTGCTCTTCCGAGGC. The use of this oligonucleotide pair and the high annealing temperature assured specificity for the P450IID6 gene. The resulting fragment of 334 bp was then digested with the restriction enzyme *Bst*N1 and the products separated on an 8% polyacrylamide gel. Bands were visualized by ultraviolet irradiation of the gel stained with ethidium bromide.

Figures 5 and 6 show the rationale and (schematically) the results.

SUBSTITUTE SHEET
ISA/EP

Examination of the dbl sequence over the intron 3-exon 4 junction suggests two explanations for the base-pair (G) deletion in the cDNA sequences. The dbl intron 3-exon 4 junction has the sequence CCCCCAG/**GACGCC** (the bold letters indicate the start of exon 4)[19]. Therefore, either a base-pair (G) deletion to give CCCCCAG/**ACGCC**, or a G to A transition to give the sequence CCCCCAAG/**ACGCC**, which shifts the position of the 3' splice site, will result in the loss of the first base (G) in exon 4. In both cases the *Bst*NI restriction site is lost. To establish which was the case we sequenced the PCR amplification product from 20 affected individuals. In all of these the G to A transition was shown to be the mutation responsible for the poor metabolizer phenotype. This transition appears to be the primary defect responsible for the poor metabolizer phenotype. Over 80% of individuals tested were homozygous for the G to A transition. Interestingly, individuals with this mutation were the same as those with the mutations at position 100 and in exon 2.

EXAMPLE 4:

Referring to Figure 7, an assay based on amplification using the oligonucleotides marked 2 and 3 followed by digestion with *Hha*I may be used to show the presence or absence of the G to A transition at bp 1094. In this case all other IID sequences will cut apart from the "b" variant. This is

SUBSTITUTE SHEET
ISA/EP

informative for the PM phenotype. Thus, the region of DNA between 1049 and 1173 is amplified, exposed to *Hha*I and submitted to gel electrophoresis followed by labelling with suitable probes. For the "b" variant, two fragments of 45 and 79 bases will be produced, whereas for EM phenotypes a single fragment of 124 bases is produced. This is an example of an indirect assay for the bp deletion at 506 in the *CYPD2D6* gene.

REFERENCES

1. Eichelbaum, M (1988) *Atlas of Science: Pharmacology*, pp. 243-251.

2. Gonzalez, F.J., Skoda, R.C., Kimura, S., Umeno, M., Zanger, U.M., Nebert, D.W., Gelboin, H.V., Hardwick, J.P. & Meyer, U.A. (1988) *Nature*, 331, 442-446.

3. Nebert, D.W., Nelson, D.R., Adesnik, M., Coon, M.J., Estabrook, R.W., Gonzalez, F.J., Guengerich, F.P., Gunsalus, I.C., Johnson, E.F., Kemper, B., Levin, W., Phillips, I.R., Sato, R. & Waterman, M.R. (1989) *DNA* 8, 1-14.

4. Skoda, R.C., Gonzalez, F.J., Demierre, A., & Meyer, U.A. (1988) *Proc. Natl. Acad. Sci. U.S.A.* 85, 5240-5243.

5. Saiki, R.K., *et al* (1988), *Science* 239, 487-91.

6. Feinberg, D.P. & Vogelstein, B. (1983) *Anal. Biochem.* 136, 6-13.

7. Kwok, S.C.M., Ledley, F.D., DiLella, A.G., Robson, K.J.H., & Woo, S.L.C. (1985) *Biochem* 24, 556-561.

SUBSTITUTE SHEET
ISA/EP

8. Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H., & Roe, B.A. (1980) *J. Mol. Biol.* 143, 161-178.

9. Biggin, M.D., Gibson, T.J., & Hong, G.F. (1983) *Proc. Natl. Acad. Sci. U.S.A.* **80** , 3963-3965.

10. Staden, R. (1986) *Nucl. Acids. Res.* 14, 217-231.

11. Kimura, S., Urmeno, M., Skoda, R.C., Meyer, U.A. and Gonzalez, F.J. (1989) *Am. J. Hu. Genet.* 45, 889-904

CLAIMS

1. A method of identifying the presence or absence of a mutation in the DNA sequence of the P450 IID bufuralol-1'-hydroxylase indicative of the EM/PM phenotype, with at least 80% certainty.

2. A method of identifying the presence or absence of a mutation at one or more of positions 100, 271, 281, 294, 408, 506 or 1432 of the DNA sequence of the P450 IID bufuralol-1'-hydroxylase or a deletion of at least part of exon 9 thereof.

3. A method according to Claim 2 wherein a mutation at position 271, 281, 294 or 506 is detected.

4. A method according to Claim 3 wherein a base-pair deletion at position 506 is detected.

5. A method according to any one of Claims 2 to 4 comprising the step of digesting DNA encoding at least part of the bufuralol-1'-hydroxylase amino acid sequence with a restriction enzyme which will cut, or will not cut, at or adjacent to one of the said positions according to whether the mutation is present.

SUBSTITUTE SHEET
ISA/EP

6. A method according to Claim 5 comprising the step of amplifying the amount of a selected region of DNA before digesting the DNA as said.

7. A method according to any one of Claims 2 to 4 wherein sample DNA is subjected to the polymerase chain reaction using oligonucleotide primers which are capable of hybridising selectively either to the wild-type or to the mutant sequence at the location being analysed, such that the generation of amplified DNA will indicate whether the said mutation is present.

8. A method of identifying a mutation in the DNA sequence of P450IID bufuralol-1'-hydroxylase at one or more of positions 100, 271, 281, 294, 408, 506, 775 or 1432 comprising (1) subjecting the sample DNA to the polymerase chain reaction using oligonucleotide primers which are capable of hybridising selectively either to the wild-type or to the mutant sequence at the location being analysed, the primers being such as to introduce, upon hybridisation to the said location, a restriction site which is not present in the wild-type or mutated location and (2) subjecting amplified DNA derived from step (1) to a restriction digest with an enzyme which cleaves at the said restriction site.

SUBSTITUTE SHEET
ISA/EP

9. A method according to Claim 8 wherein two pairs of primers are used in step (1), one pair hybridising on respective sides of position 506 and the other pair hybridising on respective sides at position 775.

10. A single-stranded DNA compound suitable for use as a primer in a polymerase chain reaction, the compound being adapted to hybridise to a region of wild-type or mutant bufuralol-1'-hydroxylase-encoding DNA flanking positions 100, 271, 281, 294, 408, 506 or 1432 such that, in the polymerase chain reaction, DNA synthesis will or will not proceed from the primer towards the said site according to whether there is a mutation at the relevant said position.

11. A compound according to Claim 10 selected from the group consisting of:

(A) 5'-TTGCGGCGCCGCTTCGGGGA-3'

(B) 5'-CTTGGGAACGCGGCCCGAAA-3'

(C) 5'-CGCCTTCGCCAACCACTCCG-3'

(D) 5'-AAATCCTGCTCTTCCGAGGC-3'

(E) 5'-GATGAGCTGCTAACTGAGCCC-3'

(F) 5'-CCGAGAGCATACTCGGGAC-3'

(G) 5'-GATGAGCTGCTAACTGAGCCC-3'

(H) 5'-CCGAGAGCATACTCGGGAC-3'

and analogues and fragments thereof.

12. A kit for performing the method of Claim 2 comprising a compound according to Claim 10 or 11.

13. A kit according to Claim 12 further comprising at least one restriction enzyme useful in distinguishing one or more of the said polymorphisms.

14. A kit according to Claim 12 further comprising a specific oligonucleotide probe useful in distinguishing one or more of the said polymorphisms.

15. A kit according to Claim 12, 13 or 14 further comprising at least one control sample of DNA containing the polymorphic site or sites and corresponding to DNA from one or more of the following: a homozygous extensive metaboliser, a homozygous poor metaboliser and a heterozygous individual.

16. A kit according to any one of Claims 12 to 15 further comprising means to facilitate obtaining cells from a patient, from which DNA may be extracted for analysis using the said method.

17. Any novel feature or combination of features disclosed herein.

SUBSTITUTE SHEET
ISA/EP

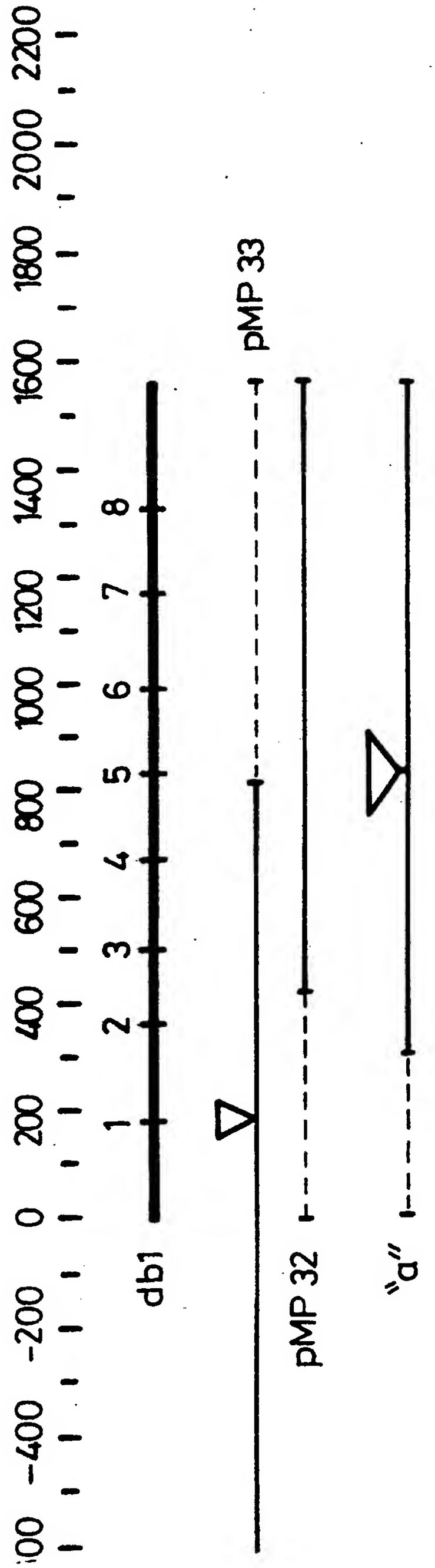
-600
-400
-200
0
200
400
600
800
1000
1200
1400
1600
1800
2000
2200
1
2
3
4
5
6
7
8
db1
pMP 32
pMP 33
"a"

Fig. 1

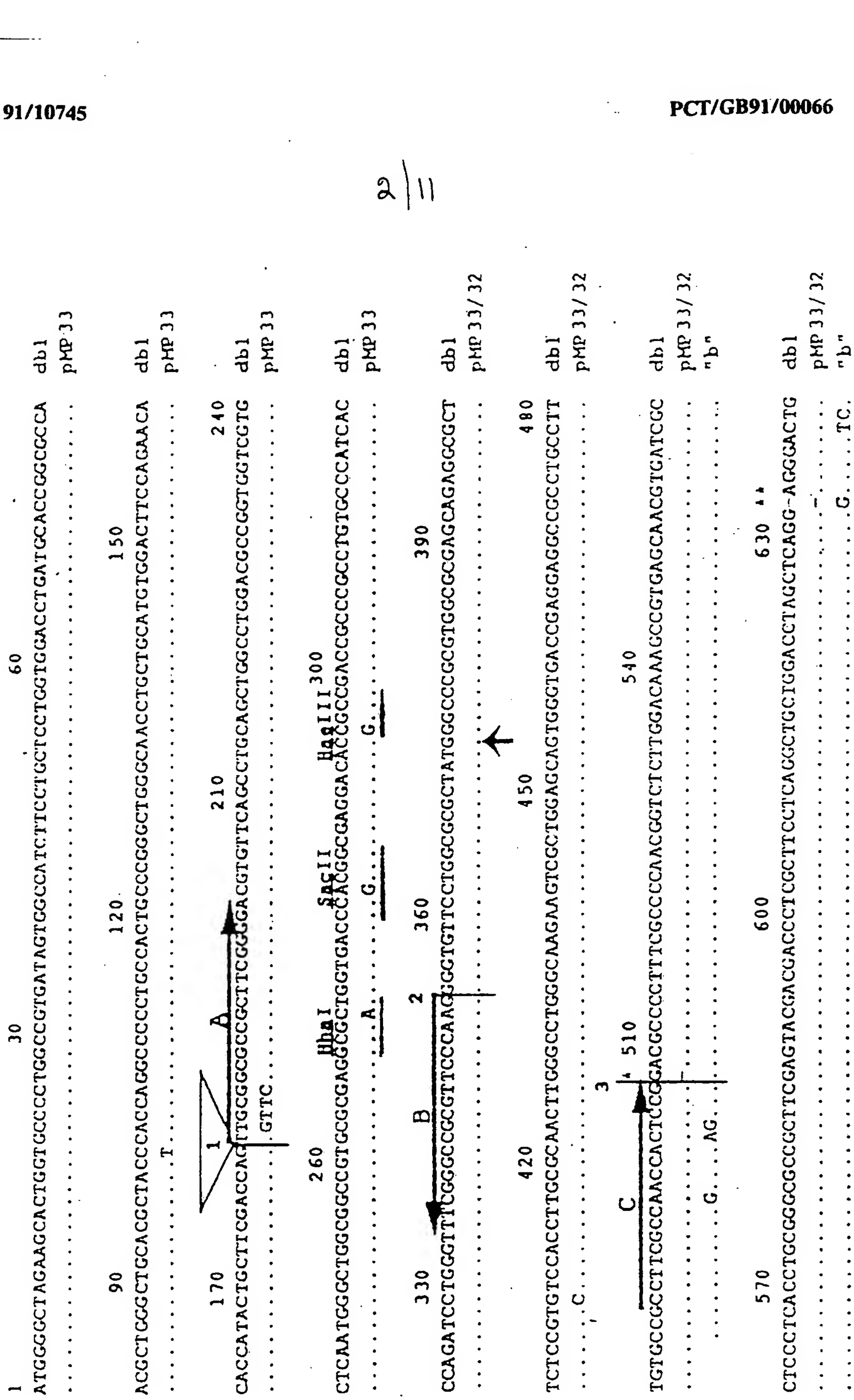

Figure 2 (1 of 3)

SUBSTITUTE SHEET
ISA/EP

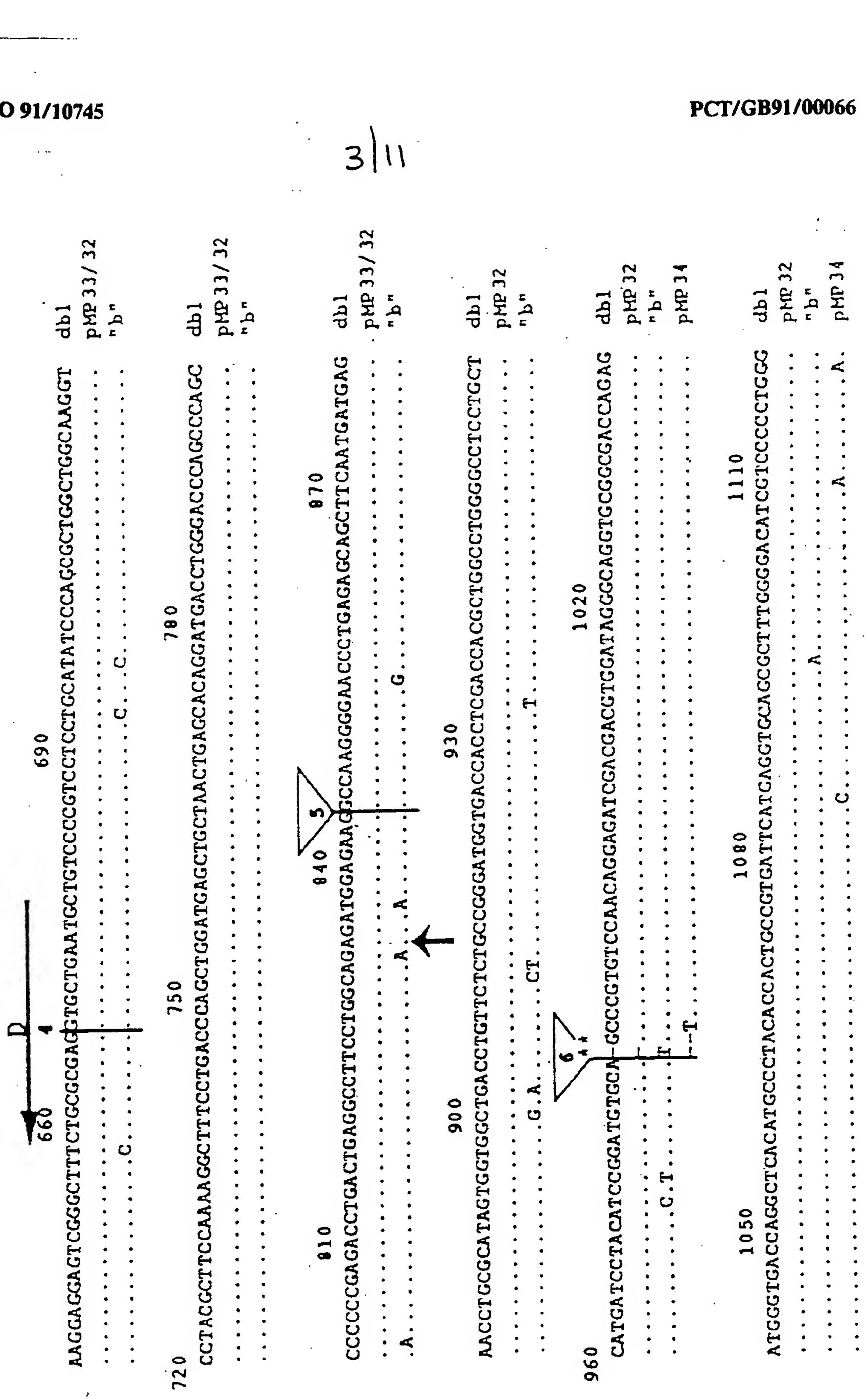

Figure 2 (2 of 3)

SUBSTITUTE SHEET
ISA/EP

```
                    1140                                                 7  1180
TATGACCCATATGACATCCCGTGACATCGAAGTACAGGGCTTCCGCATCCCTAAGGGAACGACACTCATCACCAACCTGT   db1
.G..................................................................................   pMP 32
.G..................................................................................   "b"
.G..................................................................................   pMP 34

1200                1230                                                 1260
CATCGGTGCTGAAGGATGAGGCCGTCTGGGAGAAGCCCTTCCGCTTCCACCCCGAACACTTCCTGGATGCCCAGGGCCAC   db1
....................................................................................   pMP 32
....................................................................................   "b"
.............................A......................................................   pMP 34

          1290                                                 E       1350
TTTGTGAAGCCGGAGGCCTTCCTGCCTTTCTCAGCAGGCCGCCGTGCATGCCTCGGGGAGCCCCTGGCCCGCATGGAGCT   db1
....................................................................................   pMP 32
....................................................................................   "b"
....................................................................................   pMP 34

                1380                                  HaeIII      1420          NcoI
CTTCCTCTTCTTCACCTCCCTGCTGCAGCACTTCAGCTTCTCGGTGCCCACTGGACAGCCCCGGCCCAGCCACCATGGTG   db1
....................................................................................   pMP 32
..................................C...G..G.C.....................TC.C...          "b"
..................................C...G..G.C.....................TC.C...          pMP 34

1440                BglII           1470                                 1500
TCTTT-GCTTTCCTGGTGAGCCCATCCCCCTATGAGCTTTGTGCTGTGCCCCGCTAGAATGGGGTACCTAGTCCCCAGCC   db1
.....-.............C.............................................................   pMP 32
..G.CA.....-.....C.....................................................C.....   "b"
..G.CA.....-.....C.....................................................C.....   pMP 34

                                                      F
          1430                                                                      db1
TGCTCCTAGCCCAGAGGCTCTAATGTACAATAAAGCAATGTGGTAGTTCC(A)n                              pMP 32
......CTAG..........................................                                "b"
..........T.........................................                                pMP 34
......CTAG..........................................
```

Figure 2 (3 of 3)

F/11

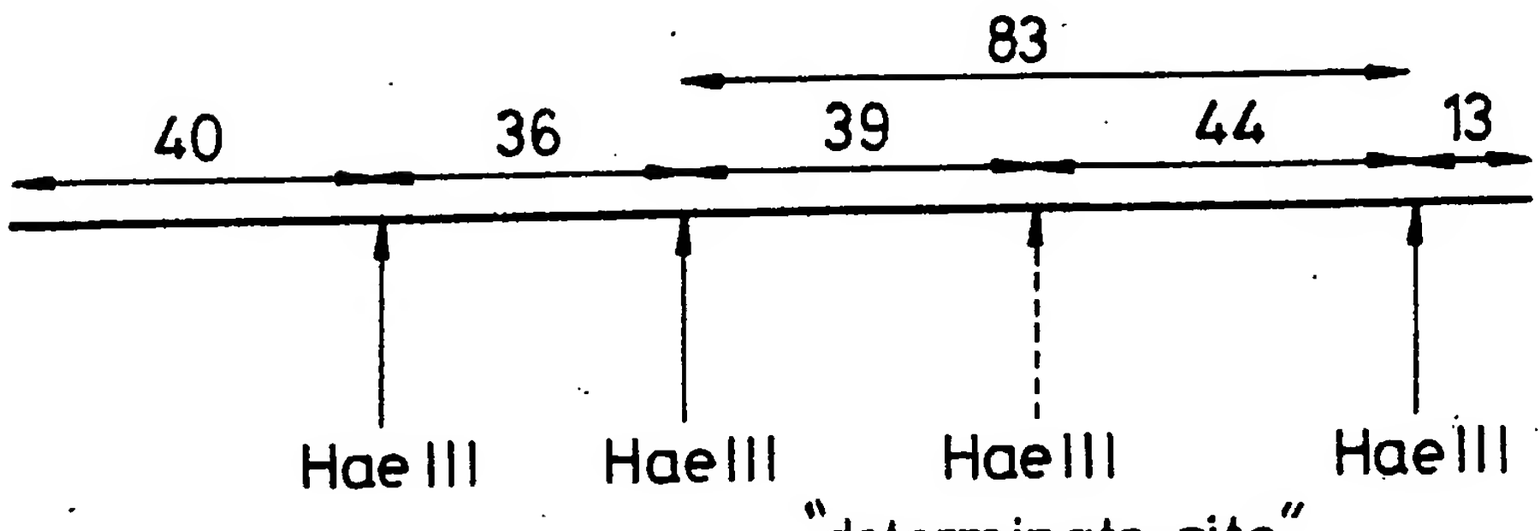
EXON 2 PCR PRODUCT DIGESTION
83
40
36
39
44
13
HaeIII
HaeIII
HaeIII
HaeIII
"determinate site"

Fig. 3

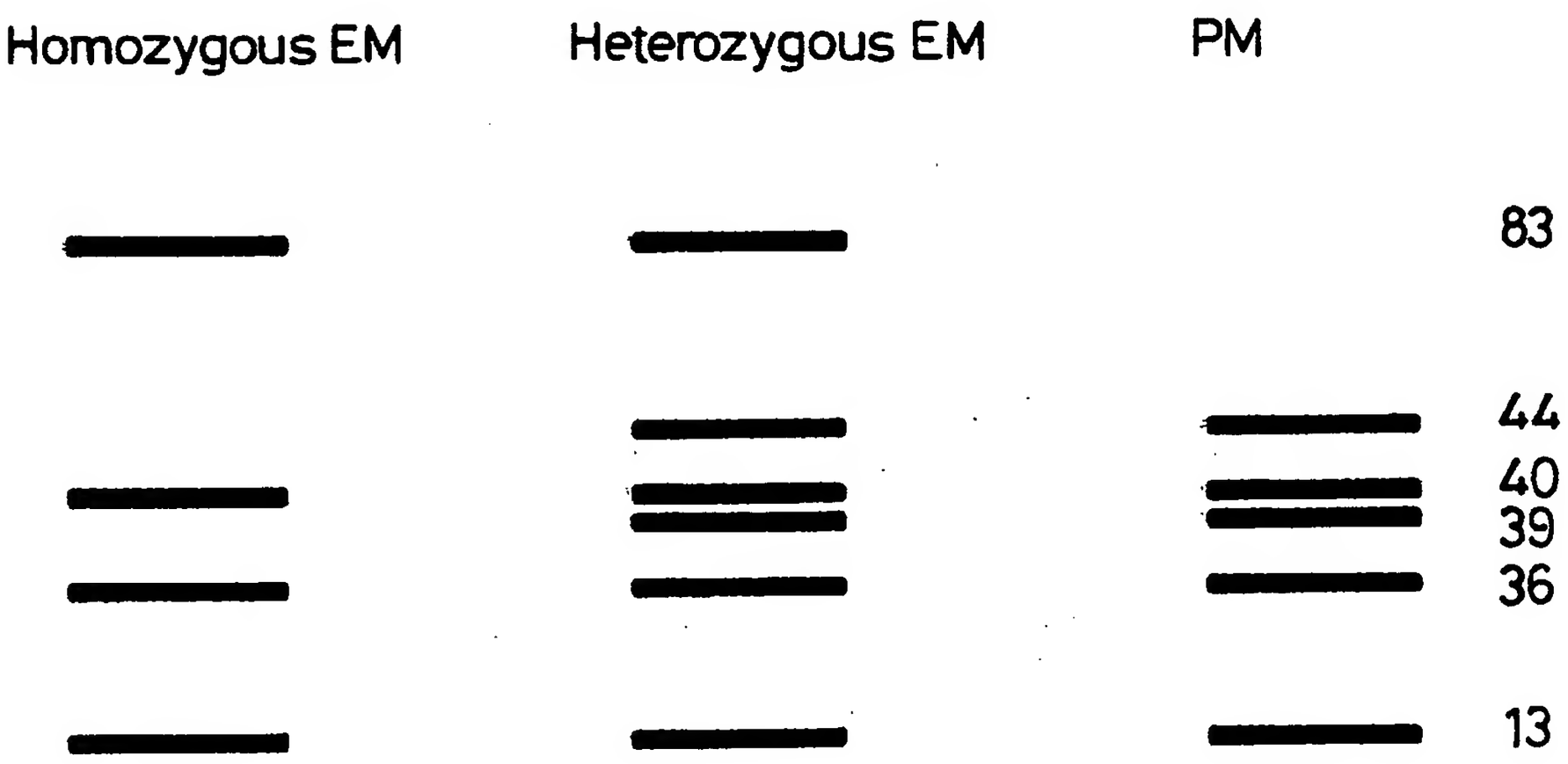
Homozygous EM
Heterozygous EM
PM
83
44
40
39
36
13

Fig. 4

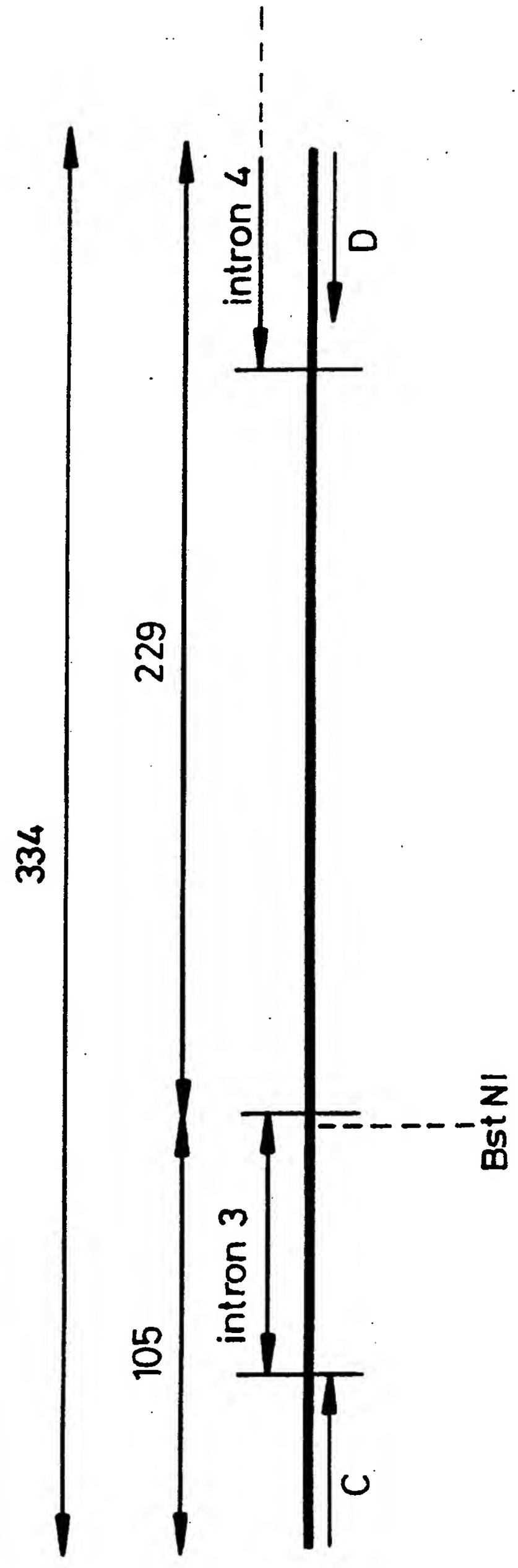

Fig. 5

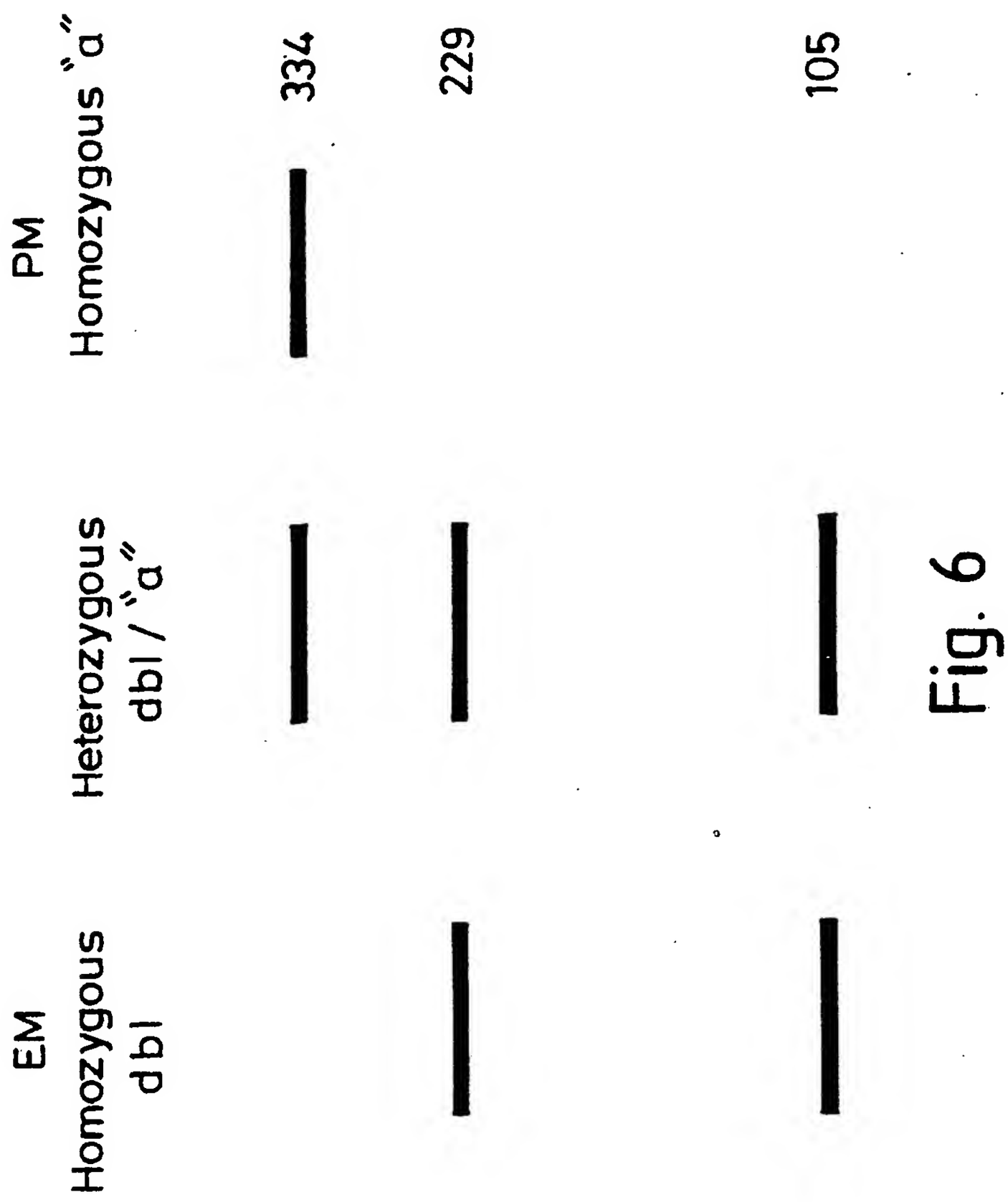
PM
Homozygous "a"
334
229
105
Heterozygous dbl / "a"
EM
Homozygous dbl

Fig. 6

```
1                              30                                  60
ATGGGGCTAGAAGCACTGGTGCCCCTGGCCGTGATAGTGGCCATCTTCCTGCTCCTGGTGGACCTGATGCACCGGCGCCA  dbl (CYP2D6)
.....................................................................................  pMP33 (CYP2D6)
.....................................................................................  CYP2D7
.............................A.....................................................A...  CYP2D8P
........G..T.................A....C.....................................G.A..A...  CYP2D8P

        90                             120                                 150
ACGCTGGGCTGCACGCTACCCACCAGGCCCCCTGCCACTGCCCGGGCTGGGCAACCT-GCTGCATGTGGACTTCCAGAACA  dbl (CYP2D6)
.....................T.............................................................  pMP33 (CYP2D6)
.....................G.....T.....................................T.....................  CYP2D7
.....A...............G.....................................T.-.....................  CYP2D8P

        170                1                210                             240
CACCATACTGCTTCGACCAGTTGCGGCGCCCGCTTCGGGGACGTGTTCAGCCTGCAGCTGGCCTGGACGCCGGTGGTCGTG  dbl (CYP2D6)
.....................GTTC.....................................................  pMP33 (CYP2D6)
.....................|.....................................................  CYP2D7
.....................|.....................................T...............  CYP2D8P
---T....AC....A.....C|.....A.....T.....................................T.....  CYP2D8P

                   260                                          300
CTCAATGGGCTGGCGGCCGTGCGCGAGGCGCTGGTGACCCACGGCGAGGACACCGCCGACCGCCCGCCTGTGCCCATCAC  dbl (CYP2D6)
.............................A.........G...........G.............................  pMP33 (CYP2D6)
.............................A.........G...........G.................C.......TA  CYP2D7
.........................T.....T.........TG.....................................C.......TA  CYP2D8P

        330                          1  360                                 390
CCAGATCCTGGGTTTCGGGCCGCGTTCCCAAGGGGTGTTCCTGGCGCGCTATGGGCCCGCGTGGCGCGAGCAGAGGCGCT  dbl (CYP2D6)
.....................................|.....................................................  pMP33/32 (CYP2D6)
....G.......C........................|.....................................................  CYP2D7
....G.......CA.........C.............|.....T.....A.A...C..A.A.....................  CYP2D8P

                   420                          450                             480
TCTCCGTGTCCACCTTGCGCAACTTGGGCCTGGGCAAGAACTCGCTGGAGCAGTGGGTGACCGAGGAGGCCGCCTGCCTT  dbl (CYP2D6)
.......C.....................................................................................  pMP33/32 (CYP2D6)
.....................................................................................  CYP2D7
.....................................C.......G.....................................C  CYP2D8P

                              3
                   510                                          540
TGTGCCGCCTTCGCCAACCACTCCGGACGCCCCTTTCGCCCCAACGGTCTCTTGGACAAACCCGTGAGCAACGTGATCGC  dbl (CYP2D6)
.....................|.....................................................  pMP33/32 (CYP2D6)
.....................|.....................................................  "b" (CYP2D9)
.........G....AG.....|.....................................................  CYP2D7
.............G....AG.....CG.....................................................  CYP2D8P
.............G....AG..A|.....A.....C...C..A.....G.C.....  CYP2D8P

                   570                          600                             630
CTCCCTCACCTGCGGGCGCCCCTTCGAGTACGACGACCCTCGCTTCCTCAGGCTGCTGGACCTAGCTCAGG-AGGGACTG  dbl (CYP2D6)
.....................................................................................-.....  pMP33/32 (CYP2D6)
.....................................................................................Ⓒ....TC.  "b" (CYP2D9)
.....................................................................................-.....  CYP2D7
.....................................................................A...........A-.....T..  CYP2D8P
.....T.....................................A.....A-.....T..  CYP2D8P
```

Figure 7

SUBSTITUTE SHEET
ISA/EP

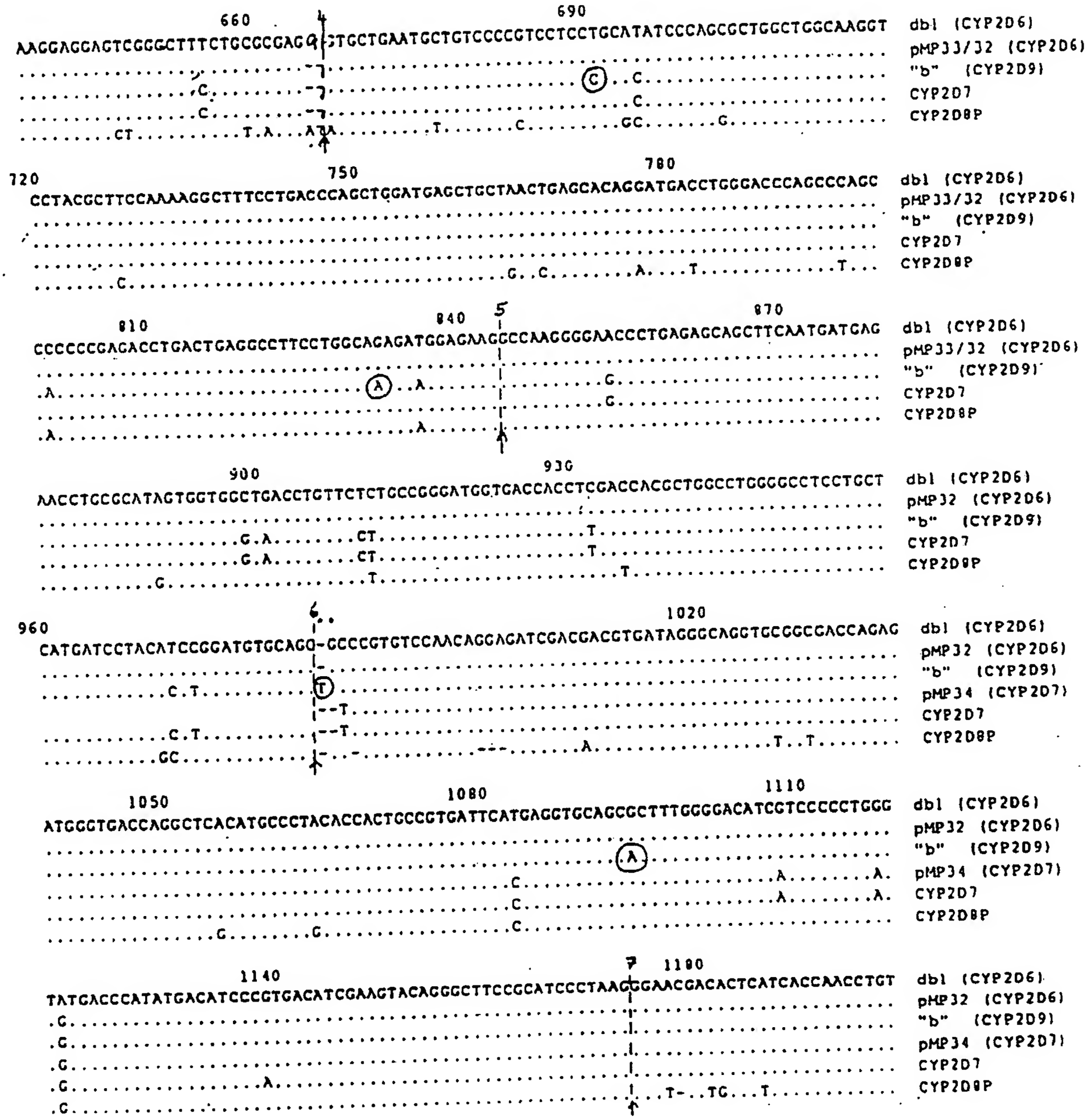

Figure 7 (continued)

SUBSTITUTE SHEET
ISA/EP

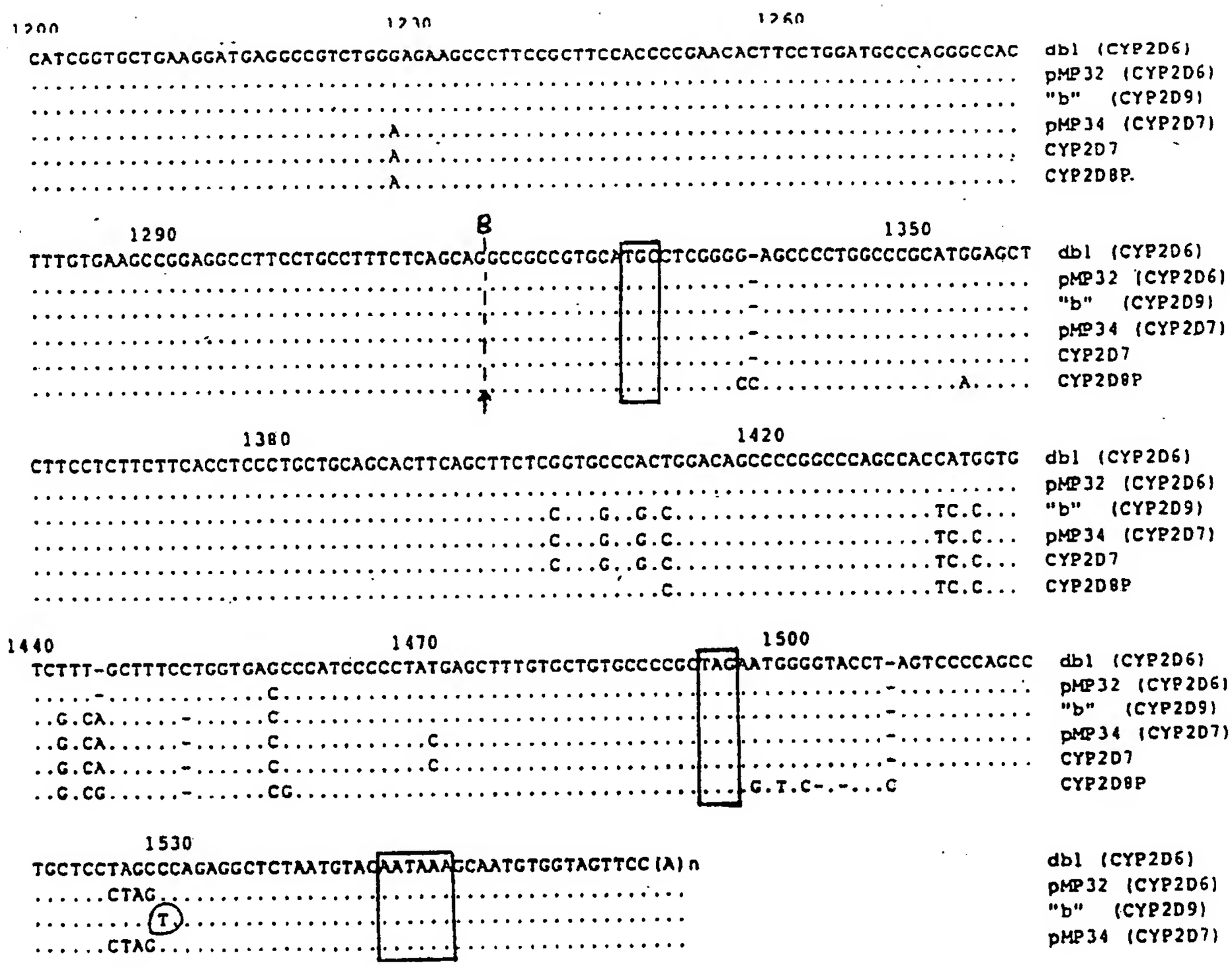
1200
1230
1260
CATCGGTGCTGAAGGATGAGGCCGTCTGGGAGAAGCCCTTCCGCTTCCACCCCGAACACTTCCTGGATGCCCAGGGCCAC dbl (CYP2D6)
pMP32 (CYP2D6)
"b" (CYP2D9)
pMP34 (CYP2D7)
CYP2D7
CYP2D8P.
1290
1350
TTTGTGAAGCCGGAGGCCTTCCTGCCTTTCTCAGCAGGCCGCCGTGCATGCCTCGGGG-AGCCCCTGGCCCCGCATGGAGCT dbl (CYP2D6)
pMP32 (CYP2D6)
"b" (CYP2D9)
pMP34 (CYP2D7)
CYP2D7
CYP2D8P
1380
1420
CTTCCTCTTCTTCACCTCCCTGCTGCAGCACTTCAGCTTCTCGGTGCCCACTGGACAGCCCCGGCCCAGCCACCATGGTG dbl (CYP2D6)
pMP32 (CYP2D6)
"b" (CYP2D9)
pMP34 (CYP2D7)
CYP2D7
CYP2D8P
1440
1470
1500
TCTTT-GCTTTCCTGGTGAGCCCATCCCCCTATGAGCTTTGTGCTGTGCCCCGCTAGAATGGGGTACCT-AGTCCCCAGCC dbl (CYP2D6)
pMP32 (CYP2D6)
"b" (CYP2D9)
pMP34 (CYP2D7)
CYP2D7
CYP2D8P
1530
TGCTCCTAGCCCAGAGGCTCTAATGTACAATAAAGCAATGTGGTAGTTCC (A) n dbl (CYP2D6)
pMP32 (CYP2D6)
"b" (CYP2D9)
pMP34 (CYP2D7)

Figure 7 (end)

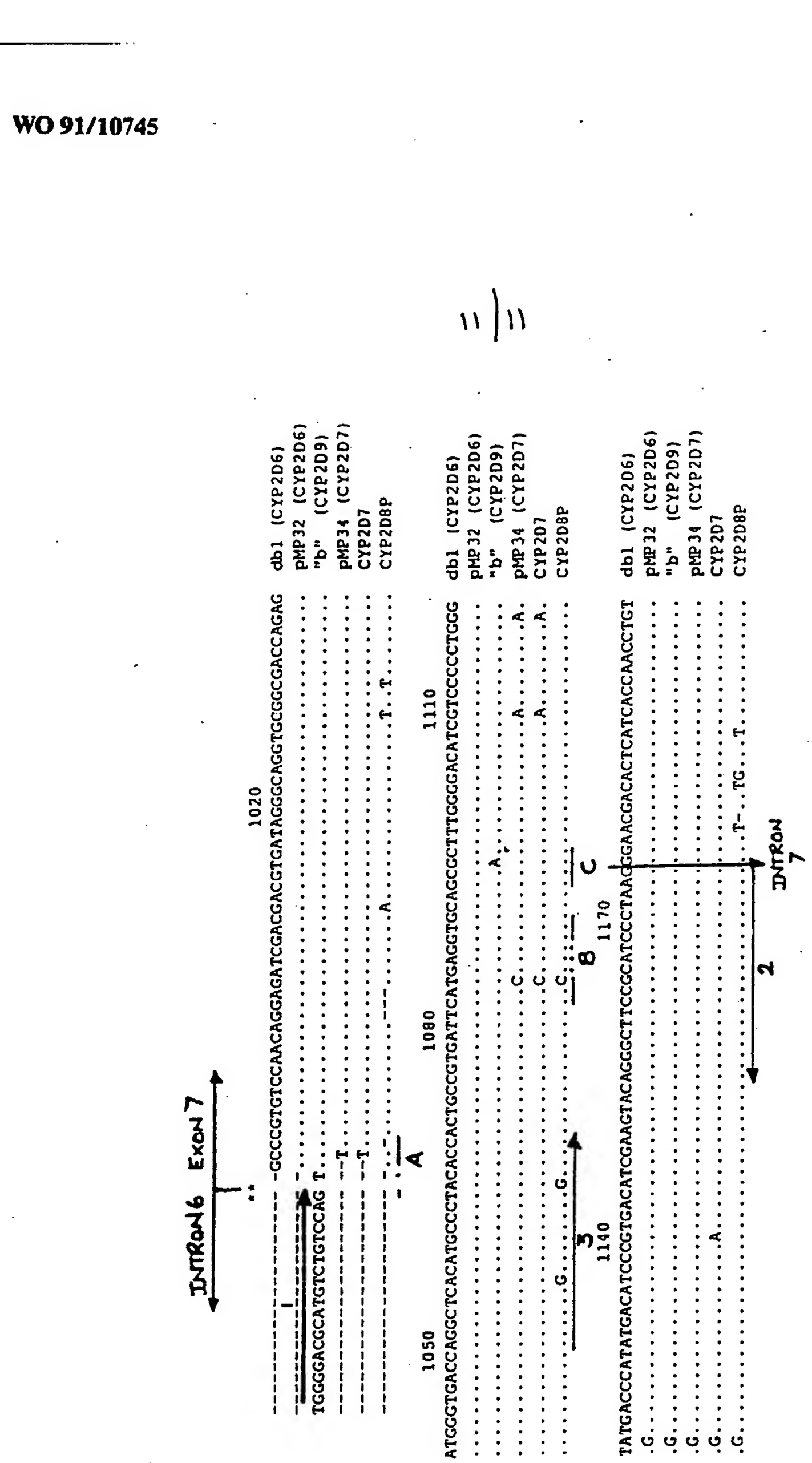
INTRON 6
EXON 7
INTRON 7
1
2
3
A
B
C
960
1020
1050
1080
1110
1140
1170
db1 (CYP2D6)
pMP32 (CYP2D6)
"b" (CYP2D9)
pMP34 (CYP2D7)
CYP2D7
CYP2D8P

Figure 8

INTERNATIONAL SEARCH REPORT

International Application No PCT/GB 91/00066

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC[5]: C 12 Q 1/68, C 07 H 21/04

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC[5]	C 12 Q

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category *	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
X	Chemical Abstracts, volume 112, 1990, (Columbus, Ohio, US), see page 266, abstract no. 71647h & US, A, 292815 (UNITED STATES DEPT. OF HEALTH AND HUMAN SERVICES), 15 July 1989 see the whole abstract ---	1
P,X	The Lancet, volume 336, 1990, M. Heim et al.: "Genotyping of poor metabolisers of debrisoquine by allele-specific PCR amplification", pages 529-532 see the whole document ---	1
A	Proc. Natl. Acad. Science, volume 85, July 1988, Medical Sciences, (US) Radek C. Skoda et al.: "Two mutant alleles of the human cytochrome P-450db1 gene (P450C2D1) associated with genetically deficient metabolism ./.	1

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
26th March 1991	16 MAY 1991
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	MISS T. TAZELAAR

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, 11 with indication, where appropriate, of the relevant passages	Relevant to Claim No.
	of debrisoquine and other drugs", pages 5240-5243 see the whole document cited in the application	

A	Science, volume 293, 29 January 1988, R.K. Saiki et al.: "Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase", pages 487-491 see the whole document cited in the application	6-9

A	Am. J. Hum. Genet., volume 45, 1989, (Chicago, US), S. Kimura et al.: "The human debrisoquine 4-hydroxylase (CYP2D) locus: Sequence and identification of the polymorphic CYP2D6 gene, a related gene, and a pseudogene", pages 889-904 see page 890, column 2, paragraph 3; figures 4,5 cited in the application	10,11

A	Nature, volume 331, 4 February 1988, F.J. Gonzalez et al.: "Characterization of the common genetic defect in humans deficient in debrisoquine metabolism", pages 442-446 see figure 4 cited in the application	10,11

A	Genomics, volume 2, 1988, Academic Press, Inc., F.J. Gonzalez et al.: "Human debrisoquine 4-hydroxylase (P450IID1): cDNA and deduced amino acid sequence and assignment of the CYP2D locus to chromosome 22", pages 174-179 see figure 1	10,11

A	J. Clin. Invest, volume 83, March 1989 The American Society for Clinical Investigation, Inc., (New York, US), M.P. Manna et al.: "Major antigen of liver kidney microsomal autoantibodies in idiopathic autoimmune hepatitis is cytochrome P450db1", pages 1066-1072 see figure 3	10,11